

▼ Abyan Ardiatama

24060120140161

Praktikum ke-1 Machine Learning

```
print("hello world")
```

```
hello world
```

```
print(10+2*5)
```

```
20
```

▼ Memuat dataset ke dalam google colab menggunakan pandas dengan memasukkan url dan nam atributnya

```
import pandas
```

```
#Memuat Dataset
```

```
url = "https://archive.ics.uci.edu/ml/machine-learning-databases/iris/iris.data"  
names = ['sepal-length', 'sepal-width', 'petal-length', 'petal-width', 'class']  
dataset = pandas.read_csv(url, names=names)
```

▼ Menampilkan dataset yang telah dimuat

```
print(dataset)
```

	sepal-length	sepal-width	petal-length	petal-width	class
0	5.1	3.5	1.4	0.2	Iris-setosa
1	4.9	3.0	1.4	0.2	Iris-setosa
2	4.7	3.2	1.3	0.2	Iris-setosa
3	4.6	3.1	1.5	0.2	Iris-setosa
4	5.0	3.6	1.4	0.2	Iris-setosa
...
145	6.7	3.0	5.2	2.3	Iris-virginica
146	6.3	2.5	5.0	1.9	Iris-virginica
147	6.5	3.0	5.2	2.0	Iris-virginica
148	6.2	3.4	5.4	2.3	Iris-virginica
149	5.9	3.0	5.1	1.8	Iris-virginica

```
[150 rows x 5 columns]
```

▼ Menampilkan dimensi dari dataset

```
#Menentukan Dimensi dari Dataset
print(dataset.shape)
```

```
(150, 5)
```

▼ Menampilkan 20 baris pertama dari dataset

```
#Melihat Isi Dataset
print(dataset.head(20))
```

	sepal-length	sepal-width	petal-length	petal-width	class
0	5.1	3.5	1.4	0.2	Iris-setosa
1	4.9	3.0	1.4	0.2	Iris-setosa
2	4.7	3.2	1.3	0.2	Iris-setosa
3	4.6	3.1	1.5	0.2	Iris-setosa
4	5.0	3.6	1.4	0.2	Iris-setosa
5	5.4	3.9	1.7	0.4	Iris-setosa
6	4.6	3.4	1.4	0.3	Iris-setosa
7	5.0	3.4	1.5	0.2	Iris-setosa
8	4.4	2.9	1.4	0.2	Iris-setosa
9	4.9	3.1	1.5	0.1	Iris-setosa
10	5.4	3.7	1.5	0.2	Iris-setosa
11	4.8	3.4	1.6	0.2	Iris-setosa
12	4.8	3.0	1.4	0.1	Iris-setosa
13	4.3	3.0	1.1	0.1	Iris-setosa
14	5.8	4.0	1.2	0.2	Iris-setosa
15	5.7	4.4	1.5	0.4	Iris-setosa
16	5.4	3.9	1.3	0.4	Iris-setosa
17	5.1	3.5	1.4	0.3	Iris-setosa
18	5.7	3.8	1.7	0.3	Iris-setosa
19	5.1	3.8	1.5	0.3	Iris-setosa

▼ Menampilkan distribusi kelas data dari dataset

```
#Distribusi Kelas Data
print(dataset.groupby('class').size)
```

```
<bound method GroupBy.size of <pandas.core.groupby.generic.DataFrameGroupBy
```

▼ Menampilkan rangkuman dataset dengan hanya menampilkan beberapa nilai dari atribut

```
#Ringkasan Statistik
print(dataset.describe)
```

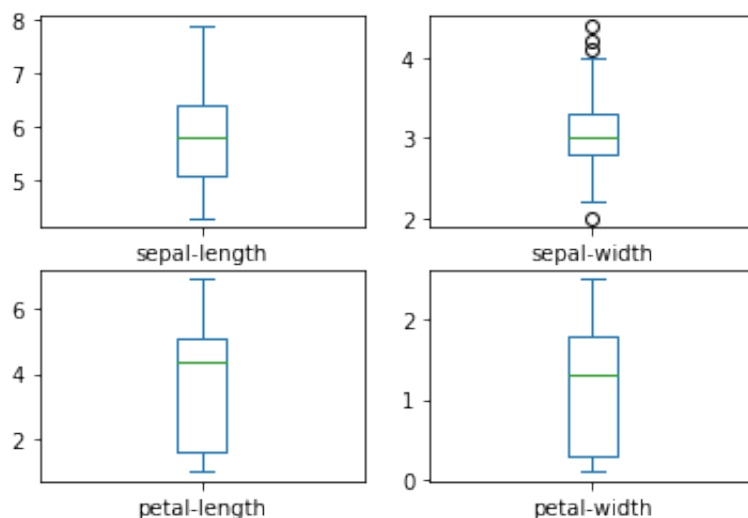
```
<bound method NDFrame.describe of      sepal-length  sepal-width  petal-len
0          5.1          3.5          1.4          0.2  Iris-setosa
1          4.9          3.0          1.4          0.2  Iris-setosa
2          4.7          3.2          1.3          0.2  Iris-setosa
3          4.6          3.1          1.5          0.2  Iris-setosa
4          5.0          3.6          1.4          0.2  Iris-setosa
..          ...          ...          ...          ...
145         6.7          3.0          5.2          2.3  Iris-virginica
146         6.3          2.5          5.0          1.9  Iris-virginica
147         6.5          3.0          5.2          2.0  Iris-virginica
148         6.2          3.4          5.4          2.3  Iris-virginica
149         5.9          3.0          5.1          1.8  Iris-virginica

[150 rows x 5 columns]>
```

▼ Memvisualisasikan dataset menggunakan matplotlib per atributnya(sepal-length, sepal-width, petal-length, petal-width)

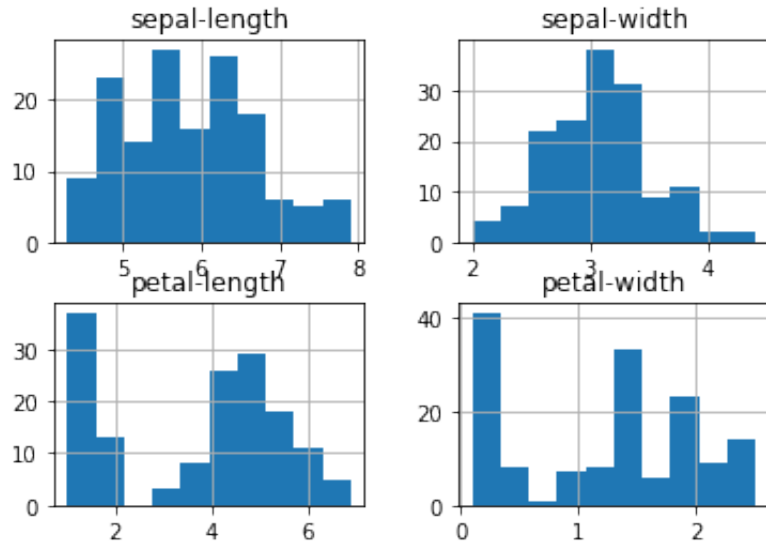
```
#Visualisasi Data Menggunakan Plot Univariat
import matplotlib.pyplot as plt
```

```
dataset.plot(kind='box', subplots=True,layout=(2,2), sharex=False, sharey=False)
plt.show()
```



▼ Menampilkan histogram dari dataset per atribut

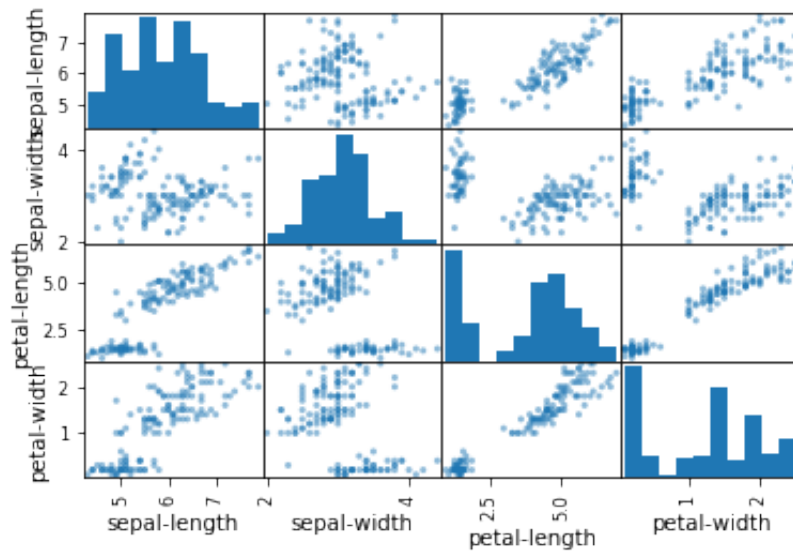
```
dataset.hist()  
plt.show()
```



Melihat interaksi antar variabel, melihat scatterplots dari semua pasang

- ▼ atribut. Hal ini dapat membantu melihat hubungan terstruktur antara variabel input.

```
#Visualisasi Data Menggunakan Plot Multivariat  
from pandas.plotting import scatter_matrix  
scatter_matrix(dataset)  
plt.show()
```



[Colab paid products](#) - [Cancel contracts here](#)

✓ 1s completed at 13:28

